Conceptual and statistical modelling of environmental effects in population dynamics

Andreas Lindén

Integrative Ecology Unit
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki
Finland

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Biosciences of the University of Helsinki for public criticism in auditorium 1041, Biocenter 2 (Viikinkaari 5) on the 5th of March 2010, at 12 noon.

Helsinki 2010
Conceptual and statistical modelling of environmental effects in population dynamics

Andreas Lindén

This thesis is based on the following articles, which are referred to in the text by their Roman numerals:


II Lindén, A., Lehikoinen, A., Hokkanen, T. & Väisänen, R. A. Modelling irruptions and population dynamics of the great spotted woodpecker — joint effects of density and cone crops. — *Manuscript*.


## Contributions

Below is a table that summarizes the contributions of all authors.

<table>
<thead>
<tr>
<th></th>
<th>Original idea</th>
<th>Material and data</th>
<th>Design and modelling</th>
<th>Analyses</th>
<th>Manuscript preparation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>LR, MF</td>
<td>LR, AL, MF</td>
<td>LR, AL, MF</td>
<td>LR, MF</td>
<td>LR, AL, VK, MF</td>
</tr>
<tr>
<td>II</td>
<td>ALe, AL</td>
<td>ALe, TH, RV</td>
<td>AL, ALe</td>
<td>AL</td>
<td>AL, ALe, RV, TH</td>
</tr>
<tr>
<td>III</td>
<td>AL, KT</td>
<td>-</td>
<td>AL</td>
<td>AL</td>
<td>AL, KT</td>
</tr>
<tr>
<td>IV</td>
<td>MP, AL</td>
<td>MP, JT, TP</td>
<td>AL, MP</td>
<td>AL</td>
<td>MP, AL, TP, JT</td>
</tr>
<tr>
<td>V</td>
<td>AL, JK</td>
<td>-</td>
<td>AL, JK</td>
<td>AL, JK</td>
<td>AL, JK</td>
</tr>
</tbody>
</table>

Mike S. Fowler (MF), Tatu Hokkanen (TH), Veijo Kaitala (VK), Jonas Knappe (JK), Aleksi Lehikoinen (ALe), Andreas Lindén (AL), Timo Pakkala (TP), Markus Piha (MP), Lasse Ruokolainen (LR), Katherine A. Thuman (KT), Juha Tiainen (JT), Risto A. Väisänen (RV)

**Supervised by**

Prof. Esa Ranta († 29.8.2008)
Prof. Veijo Kaitala
University of Helsinki
Finland

**Reviewed by**

Dr. Jan Lindström
University of Glasgow
United Kingdom

Prof. Eric Post
Pennsylvania State University
United States of America

**Examined by**

Prof. Tim Benton
University of Leeds
United Kingdom
# Table of contents

**Abstract**.................................................................................................................................6

**O  Summary**.......................................................................................................................... 7

1. Introduction.......................................................................................................................... 7

   1. 1. The interplay of intrinsic and extrinsic factors......................................................... 7

   1. 2. Appearance of the environment............................................................................. 7

   1. 3. The importance of recognizing environmental effects.......................................... 8

       Box 1: Population variability in autocorrelated environments................................. 9

   1. 4. Aims of the thesis.................................................................................................... 9

2. Environmental forcing in population dynamics.................................................................. 10

   2. 1. Mechanisms of environmental forcing................................................................. 10

       Box 2: The jungle of rate-, ratio- and density dependence........................................... 12

   2. 2. Functional form of the environment–demography relationship............................ 13

   2. 3. The spatial setting.................................................................................................... 14

       Box 3: Interaction between climate and migration..................................................... 15

3. Statistical identification of environmental effects.............................................................. 15

   3. 1. Cross-correlation can be misleading...................................................................... 15

   3. 2. Density dependent unstructured population models............................................. 16

   3. 3. Dealing with autocorrelation................................................................................. 17

       Box 4: Spurious correlations and Monte Carlo tests............................................... 18

   3. 4. Consequences of ignoring observation error......................................................... 18

       Box 5: Autocorrelation + wrong error structure = problems..................................... 19

4. Concluding remarks............................................................................................................. 20

5. Acknowledgements............................................................................................................. 21

6. References............................................................................................................................ 22

I  Ecological and evolutionary dynamics under coloured environmental variation............ 29

II Modelling irruptions and population dynamics of the great spotted woodpecker — joint effects of density and cone crops.................................................. 41

III Drought, settlement and population synchrony in North American ducks................................. 59

IV Linking weather and habitat to population dynamics of a migratory farmland songbird....................................................................................................................... 77

V Estimating environmental effects on population dynamics: consequences of observation error................................................................................................................. 95
Abstract

Population dynamics are generally viewed as the result of intrinsic (purely density dependent) and extrinsic (environmental) processes. Both components, and potential interactions between those two, have to be modelled in order to understand and predict dynamics of natural populations; a topic that is of great importance in population management and conservation. This thesis focuses on modelling environmental effects in population dynamics and how effects of potentially relevant environmental variables can be statistically identified and quantified from time series data. Chapter I presents some useful models of multiplicative environmental effects for unstructured density dependent populations. The presented models can be written as standard multiple regression models that are easy to fit to data. Chapters II–IV constitute empirical studies that statistically model environmental effects on population dynamics of several migratory bird species with different life history characteristics and migration strategies. In Chapter II, spruce cone crops are found to have a strong positive effect on the population growth of the great spotted woodpecker (*Dendrocopos major*), while cone crops of pine — another important food resource for the species — do not effectively explain population growth. The study compares rate- and ratio-dependent effects of cone availability, using state-space models that distinguish between process and observation error in the time series data. Chapter III shows how drought, in combination with settling behaviour during migration, produces asymmetric spatially synchronous patterns of population dynamics in North American ducks (genus *Anas*). Chapter IV investigates the dynamics of a Finnish population of skylark (*Alauda arvensis*), and point out effects of rainfall and habitat quality on population growth. Because the skylark time series and some of the environmental variables included show strong positive autocorrelation, the statistical significances are calculated using a Monte Carlo method, where random autocorrelated time series are generated. Chapter V is a simulation-based study, showing that ignoring observation error in analyses of population time series data can bias the estimated effects and measures of uncertainty, if the environmental variables are autocorrelated. It is concluded that the use of state-space models is an effective way to reach more accurate results. In summary, there are several biological assumptions and methodological issues that can affect the inferential outcome when estimating environmental effects from time series data, and that therefore need special attention. The functional form of the environmental effects and potential interactions between environment and population density are important to deal with. Other issues that should be considered are assumptions about density dependent regulation, modelling potential observation error, and when needed, accounting for spatial and/or temporal autocorrelation.
1. Introduction

1.1. The interplay of intrinsic and extrinsic factors

The fundamental process determining fluctuations in species abundance is referred to as population dynamics. It is a result of births, deaths, immigration and emigration. The factors affecting population dynamics through these components are commonly divided into intrinsic (purely density dependent) and extrinsic (environmental) factors, of which the latter are the focus in this thesis. After a long-lasting debate on whether intrinsic or extrinsic factors are the main driving forces in population dynamics, it has by now been widely accepted that they usually act simultaneously in producing population fluctuations (e.g., Yalden & Pearce-Higgins 1997; Forchhammer et al. 1998; Dennis & Otten 2000; Jonzén 2002a, 2005; I–V). The mutual effect of intrinsic and extrinsic factors has been referred to as the synthetic view of population regulation (Turchin 1995, 1999). The current debate is mostly about the relative roles of intrinsic and extrinsic factors in determining population dynamical patterns and the mechanisms and principles by which they affect demography.

An important finding is that intrinsic and extrinsic factors need not to be independent, but are actually likely to interact. Consequently, they have to be considered simultaneously in order to understand population fluctuations (Higgins et al. 1997; Grenfell et al. 1998; Coulson et al. 2004; Benton et al. 2006). It is therefore appropriate to distinguish between density dependent and independent environmental effects.

1.2. Appearance of the environment

The environment of an organism is an abstract concept. It includes all relevant factors surrounding the organism in focus and thus it has an unmanageable number of dimensions. To study environmental effects, one has to choose variables that approximately describe the variation in how the organism in focus experiences the environment, or some aspect of it. To include an unknown part of the environment in population models, it has to be treated as a random variable, known as environmental stochasticity.

Environmental factors can be classified in several ways according to their properties. Environment can be divided into conditions and resources (Begon et al. 1996). Conditions typically include weather and climate factors, while resources are at least in principle limited and can induce intraspecific competition. Further, the environment can be divided into the abiotic and biotic environment. Typical for the latter is that it has its own dynamics and can create second order feedback loops that affect the population in focus (Berryman 2001).

Knowledge of the basic ecology of a species is essential when choosing environmental variables to explore. Specific variables are preferable, whenever known to be relevant (Hallett et al. 2004). Such specific variables can be, for example, local temperature or air humidity measured close to the population (or individuals) in focus. On the other hand, if little is known about what the relevant factors are or the study comprises a very large geographical scale, using a large scale climate index can be advantageous (Stenseth et al. 2002, 2003). Some large scale climate indices, for example the North Atlantic Oscillation, NAO (Hurrell 1995, 1996) and the El Niño southern oscillator (Burgers 1999) are strong correlatives of many local weather variables over vast geographical ranges, often including also some relevant factors for the focal species.
Hence, they can function as proxies of how the environment is experienced. Using large scale climate indices as explanatory variables is to a large extent analogous to the use of principal components. Often large scale indices seem to better predictors of population dynamics, compared to local weather. In such cases, this can be due to poor knowledge of exactly what local weather factors properly captures the relevant variation in demography and the mechanism how weather affects demographical processes (Hallett et al. 2004). Correspondingly, a detected effect of a large scale climate variable on population dynamics, will not tell much about the exact demographical mechanisms. For example, if the average winter NAO has a positive effect on the population of a short-distance migratory bird it will not tell what aspects of weather are the relevant ones, whether they operate at the wintering quarters, during migration, once the birds have arrived to their breeding grounds, or a combination of these.

Here, and in population dynamics in general, the population and the environmental variables are temporal quantities. Data are typically time series of the relevant variables from one or several locations. A property specific for time series is that of autocorrelation (or serial correlation); the relatedness between consecutive observations (Chatfield 2004). This property is also referred to as the colour of the time series. Positive autocorrelation indicates that consecutive observations are similar in relation to the time series mean and variance. As a consequence, positively autocorrelated (or red) time series fluctuate slowly. Conversely, negatively autocorrelated (or blue) time series fluctuate rapidly, such that one extreme value is probably followed by an opposite extreme value on the other side of the average. Time series with zero autocorrelation are called white and indicates that consecutive values are in practice independent. Autocorrelation of the environment is a property that has consequences for the population dynamics (I) as well as the choice of an appropriate modelling approach when studying environmental effects (II–V).

1. 3. The importance of recognizing environmental effects

A prerequisite for assessing any consequences of environmental effects on natural populations is recognition of the relevant environmental variables. Studying environmental effects is likely to become even more pronounced due to ongoing climate change (IPCC 2007), as many species might encounter problems in adapting to the new circumstances (McCarthy 2001; Walther et al. 2002; Sæther et al. 2004) and in dispersing to new areas through highly fragmented habitats (Hanski 1999). Empirical populations are modelled for purposes such as conservation and harvesting. Sensible models of population dynamics should include environmental factors of major role in determining population demography. Both the strength and type of environmental effect influences the outcome in terms of population dynamics (Petchey et al. 1997).

In addition to explicit population dynamical modelling, recognition of the most fundamental environmental factors affecting population demography is informative as such. Investigating the properties of relevant environmental variables can tell us something about the consequences they might have on the dynamics of the focal populations. Changing means and variances of environmental factors, due to e.g. climate change, are obvious properties of interest, but also the temporal structure (autocorrelation) of the environmental variation affects the patterns of population dynamics (I).

The effect of environmental autocorrelation on population dynamics has been studied using single species models (e.g. Ripa & Lundberg 1996; Kaitala et al. 1997a, b; Petchey et al. 1997; Greenman & Benton 2003) as well as community models (Kaitala et al. 1997a; Ripa et al. 1998; Xu & Li 2003; Ruokolainen et al. 2007; Ruokolainen & Fowler 2008). According to the results of these studies, the autocorrelation of the
environment affects population variability and hence also extinction risks (Ripa & Lundberg 1996; Heino et al. 2000; Heino & Sabadell 2003; Ruokolainen et al. 2007; Ruokolainen & Fowler 2008). Although there is experimental support for these findings (Petchey 2000; Laakso et al. 2003), the exact outcome varies somewhat with the assumptions of the models. Nonetheless, general patterns of how an autocorrelated environment may affect populations are arising from recent studies as well as from the study at hand (Box 1).

**Box 1: Population variability in autocorrelated environments (I)**

In simple single-population dynamics, the variability of populations showing undercompensatory dynamics is higher in positively autocorrelated environments, and lower in negatively autocorrelated environments. Populations with overcompensatory dynamics show the opposite pattern. Hence, the population fluctuations induced by environmental variation is amplified in population dynamics with otherwise similar temporal properties as the environment (Greenman & Benton 2003).

In Chapter I, simulation of modified stochastic versions of the Ricker model suggest that the described pattern is rather general for simple systems. The figure below show some scenarios with undercompensatory (red lines) and overcompensatory dynamics (blue lines). The population variability is measured as ensemble variance, i.e. the variance of the sum of all population sizes. The pattern described for single-patch dynamics (dotted lines), can be generalized for simple spatially structured single species systems with stepping stone dispersal (solid lines) and competitive communities (dashed lines) as well.

Reprinted with kind permission from Elsevier Ltd.

1.4. Aims of the thesis

This work is inspired by the fact that the statistical identifiability of population dynamical patterns can be poor, especially without a priori knowledge about the dynamical properties of the focal population. First, dominant demographic processes can mask environmental effects of interest, if not accounted for in the modelling process (Ranta et al. 2000; Kaitala & Ranta 2001; Lundberg et al. 2002). Secondly, intrinsic and extrinsic factors can produce similar patterns of population dynamics, possibly resulting in misleading diagnoses (Jonzén et al. 2002b). These challenges, combined with
shortcomings in our knowledge about populations being modelled (e.g. short term datasets or imprecise estimates of population size) obscure the visibility of population dynamical processes (Ranta et al. 2006). Consequently, obtaining unbiased estimates of environmental effects and their uncertainty is a far from trivial task.

The aim of this thesis is to address some of the most fundamental biological assumptions and methodological issues that need to be considered when modelling environmental effects in population dynamics. My main focus is on statistical identification and quantification of environmental effects. I concentrate on simple approaches that are accessible and easy to use for the average ecologist. Most study species and populations are characterized by their own specific details that might be important for the properties of the system (Benton et al. 2006). Consequently, the applied modelling approaches should always be considered case-by-case. Chapters II–IV in this thesis are empirical case studies on different migratory bird species that all have unique ecology and migratory habits. These studies provide examples on how the approaches and results presented in Chapters I, V and this summary can be utilized in practice in ecological research.

2. Environmental forcing in population dynamics

2. 1. Mechanisms of environmental forcing

Environmental effects on population demography are referred to as environmental forcing. A fundamental step in modelling these effects is to hypothesize how environmental variation affects demographic parameters. One issue to consider is whether there are possible time lags in the effects, due to e.g., population structure or maternal effects (Benton et al. 2001, 2006; Kaitala & Ranta 2001). Typically, environmental effects are assumed to be multiplicative, altering demographic rates per capita. These are the kinds of effects in focus here (I–V), although the soundness of such an approach needs to be considered case-by-case. Another question of great importance is whether or not the environmental effects are thought to interact with population density.

As an example, the severity of winters in northern Europe has been shown to affect the survival of some bird species (Sæther et al. 2000; Francis & Saurola 2004). One possibility is that winter survival is lowered in all birds due to increased energy consumption, possibly because poor condition leads to e.g. higher risk of disease or predation. If the probability of survival is lowered by the environment with a constant probability, regardless of density, the effect can be summarized through a lowered population per capita growth rate. In unstructured mean-field models, such an effect of environment can be included by multiplying the model expectation of population density in upcoming breeding season, by \( \exp[bW_t] \), where \( W_t \) is an environmental variable describing the winter severity in year \( t \), and \( b \) is its effect size. This same principle can be applied when modelling environmental stochasticity (Brännström & Sumpter 2006, I), i.e. \( W \) is treated as an unknown random variable with a known probability density function. If the variable \( W \) is normally distributed with zero mean and variance \( \sigma^2 \) the expression can be modified to \( \exp[bW_t - (b\sigma)^2/2] \), such that the mean of the resulting log-normal distribution is one, and the expectation of the model remains independent of the term for environmental variation (Hilborn & Mangel 1997). For example in the Ricker model (Ricker 1954; Brännström & Sumpter 2005, I) the environmental component can be included in the exponent of the equation as:

\[
E[N_{t+1}] = N_t \exp\left[ r (1 - N_t/K) + bW_t - (b\sigma)^2/2 \right]
\]

Eqn 1
Here \( N_t \) is the population size (or density) at time \( t \), \( E[N_{t+1}] \) is the expectation for the following time step, \( r \) is the intrinsic growth rate and \( K \) is carrying capacity. This kind of density independent multiplicative effect of environment has in the literature also been denoted as “vertical perturbations” (Royama 1992).

On the other hand, the species in focus might be territorial in winter and increase the size of its territory as a response to increased energy demands. In such case the environmental effect is clearly affecting the strength of within-species competition and is therefore density dependent. There are several ways to model such interactions between density and environment. One approach is to replace the carrying capacity \( K \) (in Eqn 1) with a temporally variable function of the environment, e.g.

\[
K_t = K_{\text{avg}} \exp\left( b W_t - \frac{(b\sigma)^2}{2} \right), \quad \text{Eqn 2}
\]

where \( K_{\text{avg}} \) is the average carrying capacity (I).

When the environmental variable of interest is a resource or an extrinsic factor that affect the population in focus explicitly through the availability of resources, the problem can also be phrased in terms of rate- and ratio-dependence (Akçakaya et al. 1995; Berryman 2001). Especially in predator–prey dynamics rate- and ratio-dependence has been an issue of debate. In rate-dependence the absolute amount of resources available at a given time is essential, exactly like any environmental condition. This corresponds to the classic Lotka–Volterra type of models (Berryman 2001), and is originally based on analogy with molecular mass action (Akçakaya et al. 1995). This could be an appropriate model in cases where the resources are not really depleted by their consumers. In a discrete time population model, rate-dependence is technically identical to the scenario described in Eqn 1.

In ratio-dependence, population growth is thought to be related to the ratio of resources to consumer population size (Arditi & Ginzburg 1989; Berryman et al. 1995b). This ratio can either be a corrected measure of density, i.e. density per resource (Leslie 1948; Berryman 2001; Jonzén et al. 2005; Lindström et al. 2005), or alternatively the amount of resources per unit density (Arditi & Ginzburg 1989; Akçakaya et al. 1995; McCarthy 1996). The two approaches give rise to slightly different functional forms between the resources and population growth, but are similar in philosophy. According to the defenders of ratio-dependent theory, the ratio is a better approximation of the realized resource availability on an ecological time scale (Akçakaya et al. 1995). This approach is worth considering, especially in the cases of limited renewable resources (Berryman et al. 1995a) that are highly temporally variable in their availability (Lindström et al. 2005). Ratio-dependent models account for potential second order feedback loops that can be typical for biological resources, but that requires knowledge and/or modelling of the resource as well. Ratio-dependence is a kind of interaction between environment and density and is hence conceptually related to Eqn 2, but technically the variable \( W \) in Eqn 1 is replaced by the ratio of the given point in time.

An interesting special case is the Gompertz model (Gompertz 1825). Including a log-normal multiplicative environmental effect similarly to Eqn 1 (the Ricker model) gives

\[
E[N_{t+1}] = N_t \exp\left[ r (1 - \ln(N_t/K)) + b W_t - \frac{(b\sigma)^2}{2} \right], \quad \text{Eqn 3}
\]

On a log-scale the equation reduces to a first order autoregressive model. By denoting \( X_t = \ln N_t \), \( a_0 = r - r \ln K - (b\sigma)^2/2 \) and \( a_1 = 1 - r \), we get

\[
E[X_{t+1}] = a_0 + a_1 X_t + b W_t, \quad \text{Eqn 4}
\]
Because $r \ln K$ is part of the constant $a_0$, the environmental effect $bW_t$ could equally well be included in the carrying capacity as in Eqn 2. The model remains structurally exactly the same (III). However, including an environmental effect $b$ according to Eqn 2, corresponds to an effect $rb$ in Eqn 3. These findings imply that with Gompertz-type density dependence, it is not possible to separate between the proposed types of density dependent and density independent environmental effects. The good news is that applying this model relaxes the assumptions about type of environmental forcing. The bad news is that statistical detection of the exact nature of an environmental effect requires knowledge of the type of density dependence present and sometimes it is simply not identifiable.

**Box 2: The jungle of rate-, ratio- and density dependence (II)**

In the Gompertz model it is impossible to separate log-linear relationships to population per capita growth and carrying capacity. The model has still another similar feature in terms of identifiability. If the relationship with the environmental variable $W$ is assumed to be logarithmic ($W$ is replaced with $\ln W$) neither rate- nor ratio-dependence is identifiable. The part $a_1X_t + b\ln W_t$ of the model can be written also as $(a_1 + b)X_t + b\ln(W_t/N_t)$. This finding can equally well be concluded that the difference between rate- and ratio-dependence is not identifiable from regular Gompertz-type density dependence. This requires consideration when interpreting results.

In Chapter II the strength of irruptive migration events and population dynamics of the great spotted woodpecker (*Dendrocopos major*, L.) in southern Finland is statistically investigated in relation to population density and cone corps of Norwegian Spruce (*Picea abies*, L.) and Scots Pine (*Pinus sylvestris*, L.). The strength of irruptive migration events shows a negative relationship to spruce cone availability while population growth shows a positive relationship. Competing hypotheses of the functional relationship of cone availability are evaluated using information theoretical model selection (Burnham & Anderson 2002; Johnson & Omland 2004). The models are evaluated with Akaike Information Criterion (AIC), which provides an estimate of the relative distance from the truth. The relative weights of evidence in terms of Akaike weights are illustrated in the figure below for the competing hypotheses (no cones = NONE, rate-dependence = RATE, ratio-dependence = RATIO, and a logarithmic relationship = LN). For each separate analysis these weights sum up to one, similarly to probabilities. The result indicates very clearly that a logarithmic relationship with spruce cone availability give the best explanation for population dynamics assuming Gompertz-type of density dependence. While this saturating type of response is evident, it is not possible to separate rate- from ratio-dependence.
It can occasionally be hard to presume whether environmental forcing is expected to be density dependent or independent. An alternative modelling approach is to model the dynamics explicitly in terms of birth and death rates. Ripa & Lundberg (2000) describes a version of the Ricker model, where intrinsic growth rate $r$ and carrying capacity $K$ are replaced with per capita birth and death rates. In seasonal environments, for example for a migratory bird, it is clear that winter-time environmental effects on population dynamics are due to mortality (Peach et al. 1991) and that effects occurring in the breeding season are predominantly due to variation in nativity. This approach can be worth considering when appropriate, but problems with identifiability are likely to be present here as well.

2. 2. Functional form of the environment–demography relationship

Last section mainly focused on how demographic traits of populations can be affected by environment and how environmental variation might interact with density. The functional form of the relationship between environment and demography was also touched upon, since it is in some cases fundamentally linked to how environment interacts with density. In this section the issue of the functional form will be considered in more detail.

Usually, the functional relationship described by a model is assumed to apply only within a realistic range of the environmental variation. The local relationship of interest is typically approximated with a simple function, such as a linear or log-linear response. The environmental effect can, of course, be non-linear in shape also within the realistic range of variation of the variable of interest. For example, Mysterud et al. (2001) studied demographical traits of wild red deer (Cervus elaphus, L.) and domestic sheep (Ovis aries, L.) populations on the west coast of Norway. They found non-linear responses of body mass in relation to the large scale climate index NAO. The smallest body masses were found with intermediate values of NAO. Local precipitation, snow depth and in particular temperature in January showed similar types of non-linear relationships with NAO.

Variables such as temperature and pH have inevitably some optimum-type of response if the range of investigation goes from extremely low to extremely high values (Begon et al. 1996). Consequently, it is worth considering whether the optimum might lie within the range of realistic environmental fluctuations. Other ecologically important types of non-linear responses include those with a threshold, those with emphasis on extreme values (Easterling et al. 2000), and the saturating ones that are typical for the availability of resources. Especially the field of predator–prey dynamics have a long tradition of different types of saturating functional responses with different mechanistically justified assumptions (e.g., Holling 1959a, b; Griffiths 1969). It is also important to note that the scale used for the environmental variable might affect the realized functional form of the environmental effect. For example a linear effect of wind measured as metres per second, is non-linear if wind is measured on the Beaufort scale.

The shapes of environmental responses determine how environmental variation translates into biological processes (Laakso et al. 2001) and are critical for the resulting population dynamical consequences (Laakso et al. 2003, 2004). In terms of signal processing, the shape of the environmental response to a demographical parameter can be seen as a filter that changes the distribution and autocorrelation of the signal (Laakso et al. 2001). Since most population dynamical theory on environmental effects assumes linear or log-linear relationships, one useful approach is to filter the variable of interest prior to analysis, such that the expected response is shaped as desired. This of course requires knowledge of the system in focus and justification for the filter. Ignoring strong non-linearity in environmental effects will induce problems with statistical identifiability when using correlative methods.
2. 3. The spatial setting

An important extension of single population dynamics is inclusion of the spatial dimension — an aspect that improves the realism and applicability of population models. Most animal populations are connected to each other by dispersal. Further, the biotic and abiotic environments influencing population growth typically show some degree of spatial autocorrelation, i.e. populations located close to each other tend to experience similar environmental conditions. Dispersal and spatially autocorrelated environment are regarded as the most important mechanisms that potentially introduce spatial synchrony in population dynamics (Bjørnstad et al. 1999a, b; Ranta et al. 1999; Grøtan et al. 2005). Typically covariation between population densities decrease by distance, and the exact functional form of this decrease depend on the underlying synchronizing mechanism (Myers et al. 1997; Bjørnstad et al. 1999a; Ranta et al. 1999). Moreover, interactions between delayed density dependence and dispersal can give rise to travelling waves, where the peak densities move across space. Apart from the pattern of covariation by distance, spatial covariation that is independent of direction is called isotropic, whereas covariation that depends on direction is said to be anisotropic (Bjørnstad et al. 1999a).

Spatially correlated environmental forcing as a synchronizing mechanism in population dynamics is known as the Moran effect. One of the first examples for such effects was proposed by Moran (1953), in a study on the population dynamics of Canadian Lynx (Lynx canadensis, Kerr). According to Moran’s theorem, two or several populations subject to environmental forcing will show the same spatial correlation as the environmental factor in focus, given that the populations have the same density dependent structure (Moran 1953; Royama 1992, 2005; Engen & Sæther 2005). The Moran effect can be further generalized to more complex non-linear systems with spatially variable population dynamics (Engen & Sæther 2005; Sæther et al. 2007). Environment can in such cases also introduce synchronous effects, but the degree of synchrony is smaller if population dynamical parameters vary in space.

An important prerequisite for the Moran effect to act as a synchronizing force is that the environmental effect on population demography is similar over different parts of the geographical range. Many studies have shown latitudinal gradients in climate effects, e.g. in Canadian Lynx (Stenseth et al. 1999) and several large herbivores (Forchhammer et al. 2002; Post & Forchhammer 2004; Post 2005). A similar situation applies for environmental effects that vary according to habitat or altitude (Mysterud et al. 2000). The differences in effect sizes of the environment are likely to diminish the potential of the Moran effect (Engen & Sæther 2005; Sæther et al. 2007).

Multivariate spatiotemporal statistical models can help to enhance our understanding on how population dynamical mechanisms operate (Dennis et al. 1998; Jonzén et al. 2005). It is possible to account for spatial synchrony in the residuals and one possibility is to model it with a few parameters describing the pattern as a function of distance (Myers et al. 1997; Bjørnstad et al. 1999a). Because such an effect is at least partly due to environmental stochasticity, this can help in separating the variation from for example density dependent fluctuations or observation error. There are also advantages of spatial models when trying to detect or estimate the strength of environmental effects. Variables with low temporal resolution can be difficult to separate from arbitrary trends or spurious relationships. However, such variables might vary in space, which clearly introduces more information on how the environmental factor affects the population.
Box 3: Interaction between climate and migration (III)

The breeding populations of North American ducks show positive correlation with the number of ponds or other measures of wetness on the prairies in the southern parts of the species breeding grounds, whereas negative relationships has been recorded in northern parts (Krapu et al. 1983; Johnson & Grier 1988). Gradients in effects of drought have lately been further confirmed using population dynamical models (Viljugrein et al. 2005; Sæther et al. 2008). In the pintail (Anas acuta, L.), the pattern has been explained by that the ducks fly north over the prairies during dry years to more favourable breeding grounds (Derksen & Eldridge 1980; Hestbeck 1995); this drought displacement hypothesis has also been advocated by the other mentioned authors.

In Chapter III the spatial population dynamics of seven species of ducks (of genus Anas) in North America is investigated in relation to the Palmer Drought Severity Index (PDSI) using density dependent models with potentially spatially correlated residuals. For each species, sets of latitudinally and longitudinally arranged populations are compared in parallel. The focus is on the responses to PDSI in different areas and whether they are formed through drought displacement or other mechanisms. According to our results, wet conditions on prairies and grasslands attract high densities of breeding ducks, whereas the situation is the opposite in the northern parts of the study area. Local PDSI is in general a much better predictor of population dynamics than the PDSI in the areas first encountered during migration. This suggests that wet conditions are disadvantageous in the north and that the ducks are able to choose their breeding grounds more freely than expected by the hypothesis of drought displacement. Further, strong residual spatial synchrony is present in the latitudinally arranged set, but not in the longitudinal one.

In migratory birds, the pattern of settlement during migration obviously affects the spatial distribution of the breeding population. Climate variation is likely to affect the patterns of settlement, which will result in more complex patterns of spatial synchrony. In addition to the traditionally proposed mechanisms of synchrony (dispersal and the Moran effect), we suggest that the interaction of climate and migratory movements can be a strong determinant in forming patterns of spatial synchrony, especially in highly mobile animals such as migratory birds.

3. Statistical identification of environmental effects

3.1. Cross-correlation can be misleading

Detecting the presence of environmental effects in population dynamics and estimating their strength is usually approached with statistical methods, using time series data consisting of population counts and environmental variables. In most cases the time series are discrete: the population is counted, e.g., once every breeding season.

One common approach for finding correlative relationships between two discrete time series is cross-correlation, i.e. calculating correlation coefficients between the time series with a set of different time lags (Chatfield 2004). This approach has been widely used for detecting presence of external forcing in population dynamics (see e.g. Royama 1981, 1992). The expected lag is usually known on biological grounds — typically it is the focal or previous year — which reduces the problems with interpretation and correction for multiple testing. The approach as such is intuitive and can realistically describe situations where, e.g., environment affects the carrying capacity of the population in focus and the population tracks that carrying capacity with a time lag. Often the measure of population density is log-transformed prior to analysis, such that the variance is stabilized and the analysis corresponds to multiplicative effects of the environment rather than additive ones.

One problem with cross-correlation is that it does not account for the typically present density dependent fluctuations in population density. High correlations between environment and population density are restricted to a rather limited range of population
dynamics, mainly corresponding to the region of equilibrium stability. With more complex overcompensatory dynamics the correlation is poor. The result applies even for simple unstructured population models and for several different demographical traits (Ranta et al. 2000; Lundberg et al. 2002). In the case of age- and/or stage structured models the situation becomes more complicated and a correlation can be difficult to find, especially so for semelparous breeders (Kaitala & Ranta 2001). Scott & Grant (2004) criticized the findings by Kaitala & Ranta (2001), by claiming that the statistical model used does not correspond to the life history of the organism. Without doubt, both were right: the lesson learned is that simple correlative approaches that do not account for the most fundamental life history traits or potentially important intrinsic effects are likely to give poor results.

Another danger with the cross-correlation approach is that autocorrelation in the focal time series introduces strong biases in estimates of uncertainty, possibly resulting in spurious correlation. The results are unbiased only if both time series are approximately serially uncorrelated. Royama (1981, 1992) suggested that the time series involved (population density and the environmental variable) can be replaced with their first or second differences prior to cross-correlation analysis. Differencing will always shift the autocorrelation leftwards (towards negative autocorrelation) and hence the method helps only if: 1) both time series are positively autocorrelated, 2) differencing is done only until the time series are approximately serially uncorrelated. However, several better approaches are available. Problems induced by autocorrelation are presented more thoroughly in Section 3. 3.

3. 2. Density dependent unstructured population models

Recently, the method of cross-correlation has increasingly been replaced with the approach of explicitly fitting density dependent population models with environmental variables as covariates (see e.g., Berryman 2001; Jonzén et al. 2002a). With accurate time series data of population size or density, statistical fitting of the Ricker model with multiplicative environmental covariates (Eqn 1) is simple. It corresponds exactly to a multiple regression model, where the logarithmic per capita growth rate ($\ln N_{t+1} - \ln N_t$) is the response variable, whereas the population size/density in the previous year ($N_t$) and the environmental variable $W_t$ (or several of them) are covariates ($I, IV, V$). Interaction with density (e.g. ratio-dependence) can be modelled by using either $W_t N_t$, $W_t/N_t$ or $N_t/W_t$ as a covariate instead of $W_t$ ($I, II$). Similarly, Gompertz-type density dependence can be obtained by using $\ln N_t$ as a covariate instead of $N_t$ ($I, II, III, V$).

Compared to cross-correlation, this approach is more mechanistic and better in accordance with the current view of the nature of population dynamics (Turchin 1995, 1999). Even if one is focused on environment rather than the population dynamics in general, better estimates of environmental effects are obtained across a much wider region of dynamical behaviours, including strongly overcompensatory dynamics and slow approach towards the carrying capacity after a population crash or new colonization. Problems with autocorrelation still remain, but they are less severe when accounting for density dependence. In conclusion, many of the major problems with cross-correlation are diminished using this approach. Furthermore, the regression model turns the focus into quantitative estimates of model parameters (effect-sizes) and their associated uncertainties, instead of mere correlations. Of course, if qualitative correlations are preferred for some reason, the method of partial correlation is analogous to multiple regression.
As outlined in Section 2.1. and Box 2, the exact functional relationship between environment and demography, as well as possible interaction between the environment and density, can be hard to infer, especially if the type of density dependence is not known. This calls for careful consideration of the type and shape of density dependence applied, or investigation of several models using model selection. Some models can be fitted using multiple regression methodology, while others might require non-linear fitting techniques, e.g. based on maximum likelihood estimation and numerical search algorithms (Hilborn & Mangel 1997) or Bayesian statistics (Gelman et al. 2004). An alternative is to apply some model where the shape of density dependence is flexible, such as the theta-logistic (Sæther et al. 2002) or theta-Ricker model (Hanski 1999). Despite the general pattern of negative density dependence, some populations might at low densities show positive density dependence; a phenomenon known as the Allee effect (Allee 1931, Stephens et al. 1999). Such effects can as well be included in stochastic models, including the statistical models presented here (Dennis 2002).

3.3. Dealing with autocorrelation

In time series analyses measures of relatedness are affected not only by genuine causal relationships, but also by properties of the time series involved. For example, two time series with strong linear trends will always show a clear positive or negative cross-correlation, simply due to this property. In addition to trends, also autocorrelation affects the estimated strength of correlative relationships and is therefore an important property to be taken account in e.g. regression approaches dealing with time series. For example, two unrelated, but strongly positively autocorrelated time series will frequently give more extreme correlations compared to serially independent time series. The sampling distribution of their cross-correlation will have a zero mean, but its standard deviation will strongly depend on the autocorrelation of the two time series involved (Box 4). The standard t-test in correlative analyses (e.g. regression) hence applies only for serially uncorrelated time series. False correlation, detected due to violation of the assumption of serial independence is called spurious correlation (Pyper & Peterman 1998, Chatfield 2004).

There are alternative ways to handle autocorrelation, with their own pros and cons. A common approach is to adjust the degrees of freedom for autocorrelation (Post & Stenseth 1998; Pyper & Peterman 1998), since negatively or positively autocorrelated variables effectively contain less information compared to serially independent time series. However, this approach will not account for cases where time series differ considerably in autocorrelation (Box 4), leading to sample correlations closer to zero.

Another commonly applied approach is prewhitening of the environmental variable of interest, i.e. removal of autocorrelation and usage of the residuals instead (Pyper & Peterman 1998; Chatfield 2004). This can be done by fitting some kind of smooth trend or an autoregressive model. As long as the correlation between the removed part (model fit) and the prewhitened part (residuals) is zero, the estimated relationship between the variable and population dynamics will be unbiased, also when using density dependent models. Removing variation from the environmental variable of interest will, however, result in increased uncertainty of the estimated effect. For instance, a simulation study on simple correlation tests indicate lowered statistical power (Pyper & Peterman 1998).
Box 4: Spurious correlations and Monte Carlo tests (IV)

The cross-correlation between two unrelated variables (X and Y) is on average zero, but the standard deviation of the sampling distribution depends strongly on the autocorrelation of the two variables. The figure below shows how the standard deviation of the sample distribution is related to combinations of autocorrelations (in the range −0.9 to 0.9) for time series of 50 time steps. Each value is based on 10,000 simulations of a stationary autoregressive process (see Ripa & Lundberg 1996).

The problem with spurious correlation due to autocorrelation is easiest illustrated with the simple case of cross-correlation between a population time series (Y) and an environmental variable (X). However, the problem remains to some extent also in density dependent models fitted by multiple regression. The estimated statistical significance can be too optimistic or pessimistic.

In Chapter IV, environmental effects in a Finnish skylark (Alauda arvensis, L.) population are investigated, using the Ricker model fitted with multiple regression. The time series of skylark densities and habitat quality were strongly positively autocorrelated. To account for biased levels of significance a Monte Carlo simulation approach was applied. In total 100,000 random time series were simulated using an autoregressive process, such that all of them were autocorrelated similarly to the skylark time series. The original regression model was fitted to all simulated time series to get the distributions of test statistics under the null hypothesis of no relationship. Corrected levels of significance could be calculated. This specific approach is however restricted to frequentistic null hypothesis testing.

One approach that is likely to catch asymmetric patterns of parameter uncertainty, including those caused by autocorrelation, is using maximum likelihood estimation and basing measures of uncertainty on profile likelihoods (Hilborn & Mangel 1997). For example, asymmetric confidence intervals for environmental effects can be calculated using this approach. Within the framework of maximum likelihood estimation and profile likelihoods, a model where the residuals are assumed to be an autoregressive process might be preferable in cases with strongly autocorrelated population time series.

3. 4. Consequences of ignoring observation error

The regression models described in Section 3. 2. are based on the assumption that all observed fluctuations in the time series reflect real changes in population abundances. The unexplained part of the variation is called process error and consists of, for instance, unknown environmental and demographical stochasticity. Such models are called process error models. Typically in ecological studies, the true population sizes or densities are not known and only estimates of true population sizes are available. Such estimates can be
obtained from standardized monitoring schemes, censuses, proxies proportional to population density (Gibbs 2000), etc. The difference between estimated and true population size is hereafter called observation error.

Due to the simplicity of fitting process error models, observation error is often ignored. A well known consequence of ignoring modelling of observation error in autoregressive models is that at least first order coefficients are biased towards values corresponding to exactly compensatory dynamics (Knape et al. 2008). In many ecological population models this means in practice a negative bias, causing spurious or overestimated negative density dependence (Freckleton et al. 2006). However, process and observation error can be modelled simultaneously within the framework of state-space models (Harvey 1990; Durbin & Koopman 2001) and the approach is becoming increasingly common in the field of ecology (de Valpine 2003; Dennis et al. 2006; Knape 2008; Tavecchia et al. 2009).

Observation error that is not accounted for does not only give rise to bias in density dependence estimates but can also introduce bias in estimates of environmental effects. Specifically, if observation error is ignored and the environmental variable of interest is autocorrelated (V), the estimated effect may be positively or negatively biased depending on the signs of the autocorrelation of the environment and population dynamics (Box 5). The bias is small if the amount of observation error is negligible. Similar phenomena might be present if the shape of density dependence is not modelled correctly. If there is a considerable amount of observation error in the data and the environmental variable in focus is strongly positively or negatively autocorrelated, two options are application of state-space models in combination with careful consideration of the shape of density dependence, or alternatively application of process error models and prewhitening the environmental variable.

**Box 5: Autocorrelation + wrong error structure = problems (V)**

In Chapter V the role of error structures when estimating (density independent) environmental effects on population dynamics are investigated. The performances of process error, observation error and state-space models are compared, using the Ricker and the Gompertz type of density dependence.

Typically observation errors are temporally uncorrelated, and consequently, adding observation error to a time series changes its autocorrelation. In process error models, where the observation errors are ignored, the decreased autocorrelation is reflected in estimates of density dependence which become biased. If the environmental variable of interest is also autocorrelated the change in autocorrelation due to observation error will cause a bias in the environmental effect as well. For undercompensatory dynamics, effects of negatively autocorrelated environmental variables are underestimated and effects of positively autocorrelated variables are overestimated. For overcompensatory dynamics the situation is the opposite. Further, the estimated uncertainty from profile likelihoods will be slightly underestimated (too optimistic). Using observation error models (ignoring process error), rather good point estimates are obtained in all situations, but the uncertainties estimated from profile likelihoods are poor, being far too optimistic (95% confidence intervals include the true parameter in 70–90% of the cases). Good results with unbiased estimates of environmental effects and realistic confidence intervals are obtained in all situations using state-space models, and when using process error models with serially uncorrelated environmental covariates. It might be worth to note that parameter estimates are relatively accurate using process error models when observation errors are small, regardless of environmental autocorrelation.
4. Concluding remarks

In my thesis I present how environmental variables can be included in fairly simple density dependent unstructured population models, in order to describe or statistically detect their effects on population growth. Such models give more quantitative and mechanistically sensible answers compared to e.g., cross-correlation analyses, and the statistical visibility of the environmental effects is improved considerably. However, the task is far from trivial and potentially includes methodological problems. According to my results, I list some issues that need special consideration.

1) When modelling variation in demographical traits, the functional relationship to both environment and density should be carefully considered, since they in many cases are related (I–III). In particular, separating density dependent environmental effects from density independent ones can be difficult or impossible without considering the shape of density dependence. To deal with this issue in empirical studies, statistical model selection and/or models with flexible functional relationships can be helpful.

2) Autocorrelation is tricky, not only in the population time series (response variable), but also in the environment (explanatory variables). Despite the risk of bias in uncertainty of correlative relationships (IV), environmental effects will be biased if the environmental variable of interest is strongly autocorrelated and density dependence is poorly modelled, e.g., due to observation error in the population data (V). One way to circumvent these problems is prewhitening of the environmental variable, but some information is lost in that procedure. Other remedies include careful modelling of the shape of density dependence, and state-space modelling in the case of noisy population data (II, V).

3) In migratory animals, such as the bird species studied in Chapters II–IV, the migratory habits are likely to be important for understanding the population dynamics. Environmental factors can affect demography both during breeding and wintering (IV). Also the pattern of settlement when arriving to the breeding grounds (III, IV), or even the probability of migration taking place (II), can be influenced by environmental variation. Such effects can have consequences for spatial aspects of population dynamics, such as synchrony (III).

In general, factors affecting population dynamics are difficult to study due to the typically high complexity of the systems. A good model is usually a simplification of reality, which catches the most fundamental properties of a system in a simple parsimonious explanation. The findings of this thesis provide some general issues and ideas that should be useful when modelling environmental effects on population dynamics. However, important patterns of population dynamics are usually explained by some details in the ecology of the species in focus (Benton et al. 2006, II–IV). Consequently, it is of great importance to be familiar with the ecology of the focal species in order to conduct successful modelling.
5. Acknowledgements

First of all, I would like to dedicate this work to my supervisor and dear friend Esa Ranta, who unexpectedly passed away in August 2008. Esa’s way of making science has its strong footprint in many aspects of this work. The theme of my thesis is largely inspired by the so-called visibility problem (e.g. Ranta et al. 2000, 2006). Surprisingly enough, Esa is not a co-author in any of the included chapters. This is not an indication of Esa being a passive supervisor. Rather he encouraged me to freely collaborate with different people as long as I was doing sensible science. I am happy for that. Thank you Esa for letting me know you. You gave me a lot and I often still miss you.

Great thanks to Veijo Kaitala, who jumped in as a supervisor after Esa and helped me with finishing the PhD-process. I would also like to thank my custos Lotta Sundström for her patience and irreducible help with various practicalities, e.g. forms to be filled, which is not my favourite hobby. I am also grateful to my pre-examiners Jan Lindström and Eric Post about their encouraging and constructive comments on this work.

Without the input of all co-authors the content of this thesis would be considerably poorer. Thank you Mike Fowler, Tatu Hokkanen, Veijo Kaitala, Jonas Knappe, Aleks Lehikoinen, Timo Pakkala, Markus Piha, Lasse Ruokolainen, Katherine Thuman, Juha Tiainen and Risto A. Väsiänen. I have learned a lot during our collaboration and it has been a pleasure working with you all. I would like to acknowledge Kim Jaatinen, Jonas Knappe, Jouni Laakso, Aleks Lehikoinen, Markus Piha, Mika Rahikainen, Lasse Ruokolainen and Saija Sirkkä for their constructive and useful comments on different parts of this summary. Jari Kostet, thank you for letting me use your painting of the great spotted woodpecker female working on a pine cone on the cover of my thesis.

Of all collaborators and colleagues, I want to highlight my two Swedish friends from Lund University, Niclas Jonzén and Jonas Knappe. Great thanks to you for all long discussions about statistics, likelihoods, Bayesian modelling, state-space models, birds and beer. Thanks also to Prof. Per Lundberg for all inspiring moments and arrangement of practicalities during my stay in Lund. The 5 months time I have spent in Lund has been the most fruitful and instructive time during my PhD work. I also want to acknowledge Prof. Nils Chr. Stenseth and the CEES (at the University of Oslo) for all various kind of collaboration. Thanks to all members of the IKP. This has been a very stimulating working environment.

I am grateful to all my friends for their support and for the possibility of sometimes focusing on other things than my PhD-thesis. Chrilde, Martin and Torge, thank you for our long lasting friendship. Thanks Jaakko, Jukka, Lasse A. and all water rails from Laru. Watching birds with you have been great. Thanks JM for the possibility of playing with you. Keep on rockin’! Thank you Allu for all the time we have spent together — it has been very important to me, indeed.

Dear mom and dad, thank you for everything you have given me, not the least a sound and healthy environment to grow up in. You have been supportive and fair, and you have never pushed me to be anything else than myself. The fact that you are both ecologists might have had indirect effects on my otherwise completely independent choice of becoming one too. The heritable component behind this excellent trait might be less than 100%. Henri and Micke, thank you for being my favourite brothers.

Dear Sonja. Your loving care and support has been invaluable in all ways. Also when my professional life has been tardy, your support has been very important and concrete. Despite the making of this theses you might recall some grant applications with tight deadlines. Thank you so much for being there for me in all situations. I love you.
This work has been funded by the University of Helsinki, and by the Nordic Council of Ministers through the Nordic Centre of Excellence EcoClim. Implementing this work would not have been possible without this. Great thanks for that.

6. References


SUMMARY